DOCUMENT RESUME

ED 342 783 TM 017 938

TITLE	National Comparisons. The 1989 Massachusetts Educational Assessment Program.
INSTITUTION	Massachusetts State Dept. of Education, Boston.
PUB DATE	Sep 90
NOTE	26p.; For a description of test content and reporting categories, see ED 338 667.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Comparative Analysis; *Educational Assessment; Elementary Secondary Education; Grade 4; Grade 8; Grade 12; Mathematics Achievement; *National Programs; Reading Achievement; *State Programs; *Testing Programs; Test Results; United States History
IDENTIFIERS	*Massachusetts Educational Assessment Program; *National Assessment of Educational Progress; Science Achievement

ABSTRACT

This report compares information gathered from the Massachusetts Educational Assessment Program (MEAP) test administrations of 1986 and 1988 with data from the National Assessment of Educational Progress (NAEP) tests given those same years (mathematics and science in 1986 and reading and United States history in 1988). The focus is on how Massachusetts schools and students measure up to national performance. Achievement scores of Massachusetts students are well above those of students nationwide. At all grade levels (4, 8, and 12) and in all content areas, with the exception of grade-12 history, Massachusetts students had scores above the national averages. Reading was the area of greatest relative strength. Future reporting by Massachusetts and the NAEP may indicate that the advantage of Massachusetts students is not as high as it currently appears; nevertheless, the result is outstanding. The student population of Massachusetts appears fairly typical of the national population, particularly in terms of home and educational background. Results provide some evidence that teachers in the upper grades are not building enough on the gains of the earlier grades. Comparisons are summarized in five tables. Two appendixes contain a brief description of the MEAP and the adaptation of the NAEP to the MEAP. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED342783

CHAPTER 188

Massachusetts Educational Assessment Program

September 1990

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Raynolds, Jr.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The 1989 Massachusetts Educational Assessment Program:

National Comparisons

MASSACHUSETTS DEPARTMENT OF EDUCATION

TM017938

BEST COPY AVAILABLE

MASSACHUSETTS BOARD OF EDUCATION

Mr. James F. Crain, *Chairperson,* Gloucester
Dr. Raquel Bauman, *Vice Chairperson*, Holden

Mr. John J. Gould, Boston
Mr. William K. Irwin, Jr., Wilmington
Mr. Joseph E. Killory, Sandwich
Dr. Anne S. Larkin, Bedford
Dr. Richard R. Rowe, Belmont
Mr. Jonathan D. Silverman, Worcester
Ms. Theodora A. Sylvester, Springield
Mrs. Frances M. Turner, South Egremeont
Dr. Joan Wallace-Benjamin, Boston
Mr. Frederick A. Wang, Needham

Dr. Randolph W. Bromery, *ex officio Board Member, Interim Chancellor, Board of Regents of Higher Education*

Harold Raynolds, Jr., *Commissioner of Education, Secretary of the Board of Education*

OFFICE OF PLANNING, RESEARCH AND EVALUATION

Allan S. Hartman, Director of Research

M. Elizabeth Badger
Beverly Miyares
Brenda Thomas
Mary Robertson

The Massachusetts Department of Education insures equal employment/educational opportunities/affirmative action regardless of race, color, creed, national origin or sex, in compliance with Title IV and Title IX, or handicap, in compliance with section 504.

CHAPTER 188

September 1990

Massachusetts
Educational
Assessment
Program

The 1989 Massachusetts Educational Assessment Program:
National Comparisons

MASSACHUSETTS DEPARTMENT OF EDUCATION

— Table of Contents —

Chapter One:
Highlights 1

Chapter Two:
Introduction 3

Chapter Three:
Achievement 7

Chapter Four:
Questionnaires 11

Appendix A:
A Brief Description of the Massachusetts Educational Assessment Program 15

Appendix B:
A Description of the Adaptation of NAEP to the Massachusetts Educational Assessment Program 17

CHAPTER 1

Highlights of Most Important Findings

This report compares information gathered from the Massachusetts Educational Assessment Program test administrations of 1986 and 1988 with data from the National Assessment of Educational Progress tests given those same years (mathematics and science in 1986, and reading and U.S. history in 1988). Many important findings of the Massachusetts Educational Assessment Program have already been published (see the inside back cover for a list of publications). This report focuses on how Massachusetts schools and students measure up to national performance.

The achievement levels of Massachusetts students are well above those of students nationwide. At all grade levels and in all content areas, with the exception of history at grade 12, students in Massachusetts had higher scores than the national averages.

Reading was the area of greatest relative strength. National averages were reported for grades 4, 8, and 12. Massachusetts students in grade 4 scored closer to the national average at grade 8 than grade 4, and similarly, the Massachusetts average for grade 8 was closer to the national average for grade 12 than the average for their own grade. This finding must be tempered somewhat, however, by the concerns that NAEP has published regarding its results in reading. Future reporting by both Massachusetts and NAEP may reveal that the advantage of Massachusetts students is not quite as extreme as it appears to be from this study. Nonetheless, even if it were to be tempered somewhat, this result is outstanding.

Questionnaire responses indicate that the student population in Massachusetts is fairly typical of that nationally, particularly in terms of home and educational background. Their parents' educational attainment is not significantly different from national averages, and the proportion of students who speak a second language at home mirrors the national average. This fact adds an additional dimension to the interpretation of the high scores of students in Massachusetts, and speaks well for the educational system, particularly at the elementary grades. However, there is clear erosion of this advantage at the upper grades. There is some evidence that teachers at the upper grades are not building on the gains of the earlier grades by demanding enough of their students.

CHAPTER 2

Introduction

The Massachuetts Educational Assessment Program was established in 1985 under the provisions of the School Improvement Act of 1985 (Chapter 188). The purpose of the assessment program is twofold: to furnish information to improve curriculum and instruction and to compare schools and districts at the state and national levels. In the spring of 1986, virtually all students in grades 3, 7, and 11 were tested in reading, mathematics and science. In 1988, social studies was added to the program and the grade levels were changed to 4, 8, and 12 to match the grades tested nationally. Results for both administrations were reported extensively in both summary and focus reports, and in an earlier report on national comparisons. In the spring of 1990, the grade levels and subjects tested remained the same as those of 1988.

In order to produce the national comparisons reported here, the Massachusetts tests make extensive use of materials from the National Assessment of Educational Progress (NAEP). NAEP was chosen as the vehicle for national comparisons because its purposes are consistent with those of the Massachusetts assessment (i.e., the program is specifically designed to assess curriculum rather than to distinguish among individual students), and it provides accurate national data. To make comparisons with NAEP national results as valid as possible, the Massachusetts assessment selects the same grade levels, administration times, and subjects to be tested as NAEP.

While the same four content areas are tested by both programs (with the notable exception that Massachusetts tests the broader area of social studies, while NAEP limits itself to history and civics, and as a result, only a subsection of the Massachusetts social studies test can be compared with NAEP), and many of the questions used for the Massachusetts program are direct copies of those used by the NAEP, there are many differences between the two programs. One important difference is that while the Massachusetts program fully tests students in all four content areas every two years, the NAEP splits its testing into two cycles. NAEP tested students in mathematics and science in 1986, touching only lightly on reading. In 1988, NAEP focused on reading, history and civics, with only a small part of their effort devoted to mathematics and science.

The 1986 Massachusetts Educational Assessment Program: National Comparisons reported the differences between Massachusetts and the nation in grades 3, 7, and 11 in science and mathematics. This report will focus on the differences between Massachusetts and the nation in grades 4, 8, and 12 in reading and history. After the 1990 assessment, when NAEP assesses students in grades 4, 8, and 12 in science and mathematics, a new report that addresses all four content areas at all three grades will be produced.

Two appendices are included with this report. The first provides an overview of the Massachusetts Educational Assessment Program. The second describes the differences between the Massachusetts assessment and NAEP, and provides some detail on the effort made to assure that the data from the two assessments is as comparable as possible.

It is important to note, however, that the legislation mandating that the Massachusetts Assessment provide national comparisons also contained another provision that makes it impossible do so with a high degreee of precision. As noted above, NAEP was chosen as the vehicle for making national comparisons, in part because it is the only source of accurate national data. NAEP tests are administered, however, under a set of rules and conditions that are different from those of the Massachusetts Assessment. NAEP data are reported at the regional and national levels only; no local results, even at the state level, are made public. In contrast, Chapter 188 mandated that school and district results for the Massachusetts Assessment results be made public. The difference in reporting requirements has many effects, both direct and indirect.

The direct effects are obvious. The motivating factors for both examinees and test administrators are different when they know results will be reported locally. While no specific research has been done that measures the impact of this effect, it is reasonable to assume that an increase in performance, influenced by local reporting, could be as much as two or three percentage points. An example can be found in the California Assessment Program. Before the second year of that testing program, the State Department of Education decided to add a space on the answer sheet for twelfth grade students to record their names. Statewide performance rose two to four percentage points that year — the first year of improved performance at grade 12 in almost a decade. It seems likely that at least some of that gain came not from increased achievement levels, but from students taking a more personal interest in their performance on the test.

While indirect effects of local reporting are less obvious, they probably are far more numerous. For example, while it might have been possible to use the NAEP tests intact if no local reporting were done, once the decision to report local results was reached, it was necessary to revise those tests to make them appropriate for Massachusetts. Another example of the indirect effects that local reporting has on test design is the timing of the tests. NAEP administers its tests in 16-minute timed blocks that include more items than most students can complete. While this practice suffices for NAEP's purposes, the testing in Massachusetts would be seriously criticized if the same administrative methods were used. Massachusetts students are given all the time they need, within reason, to complete the test.

Another difference in the Chapter 188 mandates is the population to be tested. Chapter 188 calls for the testing of all students with limited exceptions. NAEP also tests all students, except for those "deemed unassessable by the school authorities." This means, in some cases, that students who would be tested for NAEP are excluded from the Massachusetts Assessment. However, since local reporting is an element of the Massachusetts Assessment, it is reasonable to assume that local administrators are more likely to exclude special needs and limited-English-proficient students from the Massachusetts testing than from national testing. Also, in part because of local reporting, the Massachusetts Assessment allows three weeks of make-up testing so that as many students as possible can be tested and included in the school report. Over 97 percent of students who were not special needs or limited-English-proficient were tested in the Massachusetts assessment in 1990. This is a larger percentage than for NAEP, where students who are absent from school during the more limited test administration period are not tested.

As a result of these effects, the comparisons provided in this report must be interpreted with caution. A more detailed description of the issues involved is provided in Appendix B.

During the 1986 assessment, special care was taken to note the effects of all changes, and an attempt was made to determine whether some of those changes had an inconsequential effect. One such change was to keep questions together as a block. NAEP organizes its questions in blocks, and has each student take three blocks. During the 1986 Massachusetts assessment, NAEP items were included in two ways. The first was to administer all the NAEP items in an intact block; the second was to intermix isolated NAEP questions with other questions constructed specifically for use within Massachusetts. There proved to be no significant differences between comparisons made on the "intact" items versus the "scattered" items. Consequently, for the 1988 assessment, no attempt was made to keep blocks of NAEP items intact. The advisory committees chose those they found most worthwhile to administer to students statewide, and deleted other items that they deemed not appropriate to include in the tests. Thus, the comparisons developed in this report are based upon all NAEP items that were used in Massachusetts tests for which Massachusetts could obtain national data.

A factor that proved to be significant, however, was the location of the blocks in the NAEP administration, and therefore this was given careful attention. As noted above, NAEP administers three blocks of questions to each student. Students perform significantly better on the first block. When comparisons were made, the data were kept as comparable as possible by using NAEP data only for those students who took the questions in the first block (thereby creating the highest possible national standard). For the 1986 assessment, the only questions used were those contained in the first testing session in the Massachusetts administration. As a consequence, data were discarded for any Massachusetts questions that were not administered during the first session. For the 1988 study, however, a few questions were used that were included in later sections of the Massachusetts tests. In order to keep the national standard as rigorous as possible, however, NAEP data were used only when students took questions during the first block of testing.

Between 1986 and 1988, many significant changes took place in the operation of NAEP. Not the least of these changes was a recognition by the U.S. Department of Education that it would be necessary to keep a larger portion of the National Assessment item pool secure. As a consequence, data for only a minimal number of items in civics and geography have been released. Many questions from these two areas were used on the Massachusetts tests, but it is not yet possible to identify which NAEP data should be matched to the Massachusetts data. As a consequence, the only area in social studies for which results are included in this report is history; reports on the other two areas will be made sometime in the future, when the data for those areas are identified.

Finally, a major problem in the interpretation of the 1986 NAEP data was created by their decision to sharply limit the amount of time students had to answer questions. As a result, large percentages of students failed to complete the tests, and the non-response rate was high for a substantial portion of their test questions. In contrast, Massachusetts permits students to take all the time they need to complete the test; therefore, virtually all questions are attempted by almost all the students. Because of this difference in administration procedures in 1986, the NAEP data were carefully reviewed before being used. Once the non-response rate for a section of a test reached 5 percent, the data from all remaining questions within that block were discarded. Consequently, the data from large numbers of questions were judged non-comparable, and the information discarded. This reduced significantly the number of questions that could be used to compare Massachusetts with the nation. For 1988, however, NAEP changed its administration procedures, enabling most students

to complete each block within the allotted time. As a result, no questions were discarded from the analysis because of non-response rate.

CHAPTER 3

Achievement

As noted in Chapter 2 and Appendix B, there are several differences between the Massachusetts Assessment and the NAEP testing programs. While there are hundreds of test questions at each grade level used by both assessments, only a limited number of questions, and only a subset of the data available on each question, can be considered reasonably comparable. Thus, while many cautions must be considered in the interpretation of the data, as noted in Appendix B, the information reported in this chapter is a carefully selected subset that can be most fairly judged comparable.

Two sets of overall results are provided in Table 1. The first is taken from the 1986 assessment, when students were tested in science and mathematics in grades 3, 7, and 11. The second set is taken from the 1988 assessment, when students were tested in reading and history in grades 4, 8, and 12.

Table 1

NAEP and MASSACHUSETTS ASSESSMENT:
Percentage Differences in Correct Responses

Grade (Year)	Content Area	Number of Questions with Comparable Data	Average Percentage of Correct Responses in Massachusetts over NAEP
3 (1986)	Mathematics	29	7.5
	Science	25	7.9
4 (1988)	Reading	20	10.1
	History	21	5.4
7 (1986)	Mathematics	36	8.3
	Science	63	5.8
8 (1988)	Reading	25	9.4
	History	29	4.9
11 (1986)	Mathematics	74	4.3
	Science	90	2.6
12 (1988)	Reading	27	7.0
	History	30	-0.8

As Table 1 shows, students in Massachusetts scored higher than their counterparts nationwide at every grade in all content areas, with the exception of history at grade 12. Students in Massachusetts scored highest relative to national averages in reading and lowest in history. Relative to national averages, students scored highest in grade 4 and lowest in grade 12; the only exception to this pattern was in mathematics, where students scored slightly higher at the middle school level than they did at the elementary level.

One possibility suggested for the lower relative test scores of Massachusetts high school students was that they might not be as relatively an advantaged group as students in the earlier grades. For example, if Massachusetts had a lower drop-out rate than the national average, it might be reasonable to explain lower test scores in light of the fact that more weak students were remaining in school in Massachusetts. However, as will be noted in the next chapter, this does not appear to be a satisfactory explanation. Students in high school in Massachusetts are no more or less advantaged than younger students. The explanation for lower scores would appear to lie elsewhere.

Table 2

NAEP AND MASSACHUSETTS ASSESSMENT:
Percentages of Questions on which Massachusetts Outscored NAEP

Grade (Year)	Content Area	Number of Questions with Comparable Data	Number of Questions on which Students in Massachusetts Outscored National Averages	Percentage of Questions on which Students in Massachusetts Outscored National Averages
3 (1986)	Mathematics	29	26	90
	Science	25	25	100
4 (1988)	Reading	20	19	95
	History	21	20	95
7 (1986)	Mathematics	36	33	92
	Science	63	56	89
8 (1988)	Reading	25	25	100
	History	29	23	79
11 (1986)	Mathematics	74	61	82
	Science	90	63	70
12 (1988)	Reading	27	27	100
	History	30	14	47

Another way of showing the extent to which students in Massachusetts outscored national averages is to look at the number and percentage of test questions on which they scored higher than the national sample. These results are shown in Table 2. Again, with the exception of grade 12 history, students in Massachusetts outscored their peers nationwide on most of the questions that were asked.

NAEP reports its data as scaled scores, which provides one final way of seeing the extent to which students in Massachusetts outscored national averages. Scaled scores are transformations of raw data that permit comparisons of groups and subgroups that are more statistically valid than those that can be made from the raw data alone. Through the use of scaled scores, groups that have taken somewhat different sets of tests questions, such as students tested in two different years, can be more fairly compared.

Computation of Massachusetts scaled scores on the NAEP scale is a complex statistical process which the contractor for the Massachusetts assessment is in the process of completing. The figures presented in Table 3 are estimates based upon simplified procedures; however, these simplified procedures generally provide accurate estimates of the same numbers that would be developed from the more complex procedures. Final linking of the Massachusetts data into the NAEP scale is being completed in time for the reporting of the 1990 results, when data for all four content areas will be available at grades 4, 8, and 12. At that time, updates of these estimates for 1988 will be reported, along with the 1990 results for both Massachusetts and the nation. Table 3 is similar to Table 1 in that both tables are based upon data from the same questions, but rather than examining the average percentage of correct responses, the reported statistic is the average scaled score.

From these data it can be seen easily just how high the scores of Massachusetts students were, especially in reading. In that area, the average score for fourth graders in Massachusetts was about at the halfway point between the national scores at grade 4 and grade 8, and the score for eighth graders was closer to the national score for twelfth graders. This is a truly remarkable result, tempered somewhat by the fact that NAEP has reported some concerns about the way it has estimated performance of students in reading. The performance of students in Massachusetts in reading is so far in excess of national averages that it may be appropriate to question the accuracy of those national averages, especially given the questions raised about them by NAEP itself. Nonetheless, the same patterns generated by the data presented in previous tables apply here: the performance of students at the elementary grades was very strong; and while students at the high school level outscored the national average in every area but history, the gap between the Massachusetts and national averages was considerably less than it was for the earlier grades.

Table 3

NAEP AND MASSACHUSETTS ASSESSMENT:
Estimated Scaled Scores

Grade	Content Area	National Average	Estimated Massachusetts Average
3	Mathematics	212.1	225.7
(1986)	Science	212.1	230.0
4	Reading	230.4	247.4
(1988)	History	220.6	229.6
7	Mathematics	267.1	278.3
(1986)	Science	248.7	263.1
8	Reading	262.8	279.6
(1988)	History	263.9	272.4
11	Mathematics	304.0	310.0
(1986)	Science	291.0	304.0
12	Reading	287.1	300.7
(1988)	History	295.0	293.7

CHAPTER 4

Questionnaire Results

In addition to the extensive sets of achievement questions that were asked of students in both the Massachusetts and the national assessments, students completed questionnaires that dealt with student background and attitudes. Additional questionnaires were developed for principals and teachers in both assessments. Several questions used in the student and principal questionnaires were the same for both the Massachusetts and national assessments, providing dimension for comparison. Several results of these questionnaires were included in the 1986 report. Most notable among the findings reported there was the unusually large percentage of schools in Massachusetts in which at least half the faculty has been there for at least 10 years. The reader is referred to that report for additional information. This chapter will review one finding from the 1986 report — the information on parent education and second languages spoken at home. The remainder of the chapter will present some additional insights gathered from the student questionnaires administered in 1988.

Review of Results from the 1986 Assessment

Information on several background factors that are known to correlate significantly with student achievement was included in the 1986 report. The 1988 NAEP data on many of these variables have not been released yet in a way directly comparable to the Massachusetts data. However, because of the importance of these variables, and because it can be reasonably assumed that the results from the 1986 assessment would be largely duplicated by the one done in 1988, the findings from the earlier assessment are reviewed here.

The results of the 1986 assessment showed that students in Massachusetts were somewhat advantaged when compared with their peers nationwide. For example, 61 percent of the students in Massachusetts at grade 3 reported that they had attended nursery school or Head Start before entering kindergarten, as compared with 55 percent of the national sample.

Similarly, several of the background questions asked at grades 7 and 11 in 1986 revealed that students in Massachusetts had home environments that were, on the average, much like those found across the country, although they were somewhat more advantaged in selected ways. At both grades, students were asked about the level of their parents' education. Nationally, 36 percent of the eighth grade students indicated that their mothers had some form of education beyond high school, compared with 42 percent of the students in Massachusetts. Similar statistics on father's education were 37 percent for the national sample and 44 percent for Massachusetts. Massachusetts students were less likely, however, to know their parents' educational levels. Twenty-two percent of the seventh grade students in Massachusetts did not know their mother's educational level, in comparison with the 17 percent of the national sample. For paternal education, the figures were 28 and 25 percent respectively.

The questions about parent education showed similar results about the match between students in Massachusetts and their national counterparts in the eleventh grade. Almost equal percentages of students in Massachusetts and in the national sample answered "I don't know" to the questions about both mother's and father's education, so the data were more clearly comparable. Forty-eight percent of the twelfth graders in Massachusetts indicated that their mothers had some education after high school, compared with 43 percent of the national sample. Similar statistics for the two groups on father's education were 53 and 48 percent respectively. Thus, the parents of students in Massachusetts were somewhat more educated, but there was not a large difference between the groups.

There was also no difference at either grade level between the responses of the Massachusetts students and the national sample to the question about foreign languages spoken in the home. While Massachusetts may be thought of as a state where second languages have a significant impact, questionnaire results indicated that the use of languages other than English was not more extensive than in other areas of the country. Whatever degree of bilingualism exists in Massachusetts is almost identical to the national cross-section.

Findings from the 1988 Assessment

Although it is clear from the 1986 data that students in Massachusetts are somewhat advantaged when compared with their peers nationwide, it was somewhat surprising to see the extent of differences between students in Massachusetts and national averages to many of the questions asked of the fourth grade students in 1988. Those results are given in Table 4.

The questions listed in Table 4 actually were asked as multiple choice questions. For ease of reporting, they were converted to appear as if they had been asked as "yes/no" questions. In fact, the percentages given reflect the proportions of students who selected specified answers.

The questions that were chosen for display in the table, and the options that were selected for reporting, were used because of their known relationship to effective performance in reading. For example, while large numbers of students watch a significant amount of television each day, both the Massachusetts and the national data show that declining test scores are associated only with excessive amounts (6 hours or more per day) of viewing. Students who watch amounts that might be considered high by adult standards (3 to 5 hours per day) actually outscored students who watched less television. Thus, the specific information chosen on which to report was the percentage of students watching 6 or more hours per day.

On all these questions, students in Massachusetts demonstrate they are being given a clear advantage over their national counterparts. Two of the questions (amount of television watching and number of books owned at home) show the commitment that parents in Massachusetts are making to ensure that the home environment is conducive to education, and the last three questions (having homework to do, and getting it done; reading significant numbers of pages of material each week; and providing time during the school day on a daily basis for the reading of books of personal choice) show that teachers are reinforcing these practices in school. Thus, it comes as little surprise to see students at grades 3 and 4 outscoring their counterparts nationwide as substantially as they do.

Table 4

NAEP AND MASSACHUSETTS ASSESSMENT:
Percentages of Students Responding Affirmatively to Selected Questions, Grade 4, 1988

Question	Percentage of Students Answering Affirmatively	
	Massachusetts	Nation
Do you watch less than 6 hours of television daily?	83	73
Do you have 30 or more of your own books at home?	65	44
Do you usually have a homework assignment each day?	93	83
Do you read more than 10 pages a day in school and for homework?	65	53
Does your teacher ask you to read books you choose yourself on a daily basis in school?	48	32

However, the significant advantage that students in the elementary grades enjoy begins to evaporate by middle school, and almost entirely disappears by the time students reach high school. Table 5 provides data similar to that of Table 4, but for grades 8 and 12.

As the two tables show, whereas 21 percent more students in grade 4 in Massachusetts than those nationwide said they have 30 or more books of their own at home, the gap shrank to 17 points in grade 8 and just 3 points in grade 12. Also, the 10 percentage point gap between television watching in Massachusetts and the nation for fourth graders shrank to 2 and 3 percentage points for students in grades 8 and 12, respectively. Similar reductions were evident for the remaining questions.

Two questions that were asked of students in grade 12 shed additional light on this issue. Students were asked, "To what extent does your English teacher expect you to discuss and analyze what you read?" Nationwide, 65 percent of students selected "a lot" as their response to this question, whereas only 53 percent of students in Massachusetts so answered. Similarly, students were asked how often their English teacher asked them to write about what they read for class. Nationwide, 55 percent of students answered either "almost every day" or "once or twice a week." In Massachusetts, those options were selected by 47 percent of the students.

Table 5

NAEP AND MASSACHUSETTS ASSESSMENT:
Percentages of Students Responding
Affirmatively to Selected Questions, Grades 8 and 12, 1988

Question	Percentage of Students Answering Affirmatively			
	Grade 8		Grade 12	
	Massa-chusetts	Nation	Massa-chusetts	Nation
Do you watch less than 6 hours of television daily?	84	82	95	92
Do you have 30 or more or your own books at home?	49	32	34	31
Does your family get a newspaper regularly?	82	77	85	83
Do you buy or get any magazines at home just for yourself?	75	65	71	68
Do you read more than 10 pages a day in school and for homework?	42	38	49	44

These last two questions seem to indicate that, at least in their English classes, less is being asked of twelfth grade students in Massachusetts than of their national peers. This is true despite the fact, at least as shown by parent education levels, students in Massachusetts are somewhat advantaged. Combined with the finding that relatively high scores are attained by students in the elementary and middle grades, but not at the high school level, it would seem that teachers of high school students in Massachusetts should be providing their students with greater challenges than they are doing currently.

APPENDIX A

A Brief Description of the Massachusetts Educational Assessment Program

In the spring of 1985, Massachusetts Department of Education staff created four advisory committees of Massachusetts teachers to help select and identify test materials to be used in the newly-created assessment program. Committees were formed for each content area (reading, mathematics, and science) to guide development of test content; a fourth committee's responsibility was to review all materials to ensure that they were fair and relevant to students of all ethnic, racial, and cultural backgrounds. In 1987, when social studies was added to the assessment, a committee was formed in that subject area.

As part of the 1988 assessment, a sample of students at all grade levels also responded to open-ended, or short answer, tests in each subject area during one of the two (at eighth and twelfth grades) or four (at fourth grade) test sessions. In 1988, a sample of these questions were scored statewide and provided the information for four focus reports published by the Department in 1989. In 1990, all open-ended questions were scored by Massachusetts teachers at two two-day scoring sessions; these results are included in the reports provided to schools and districts.

When preparing for the biennial rounds of assessment testing, the committees meet regularly to review the tests used in previous years, identify items that should be revised or replaced, and guide test development staff in the creation of new items. NAEP objectives and test questions are also reviewed for their inclusion in the tests. As part of the development process, an objectives survey is sent to all schools and districts so that frameworks for the tests reflect the current state of curriculum and teaching in Massachusetts. The committees are responsible for approving final versions of the tests.

Test administration workshops for the different grade levels are conducted throughout the state just prior to the scheduled testing dates. (Grade 12 students are tested at the end of January, and grades 4 and 8 students are tested in late March.) Test materials are delivered to the schools during the week before test administration to ensure the security of test questions.

The Massachusetts Educational Assessment Program tests are comprehensive, covering a broad range of objectives in reading, mathematics, science, and social studies. The time required to take the full test battery if individual student test scores were produced would be between eight and twelve hours, depending on the grade level. Because individual student scores are not required, however, it is not necessary to give each student the full test battery. Instead, the hundreds of test questions are divided among several forms of the test at each grade level, and each student completes only one form of the test. This procedure for testing, called matrix sampling, results in greater reliability and validity than does conventional testing, while keeping testing time to a minimum.

At each grade, students complete a common questionnaire (taken by all students) and one test form. The test forms at grades 8 and 12 are made up of two testing sessions of two subjects each;

the test form at grade 4 is composed of four sessions, one for each subject assessed. Each test session includes some NAEP items on which the national comparisons made in this report are based, but students do not take the Massachusetts Educational Assessment Program tests in a manner paralleling NAEP administration. Across the three grades, the entire battery consists of 240 questionnaire items and 2,156 test questions, about two-thirds of which are taken from NAEP.

Chapter 188 requires the testing of all students at the selected grade levels, with the exception of some special needs and limited-English-proficient students. Special needs students are exempted from the assessment program if so waived by their parents; limited-English-proficient students are exempted unless their parents request otherwise. Across the three grade levels tested, just over half of the students in both of those categories were tested, and are included in this report. Of the students not exempt from testing for either of these reasons, over 97 percent were completely tested and are included in this report.

APPENDIX B

A Description of the Adaptation of NAEP to the Massachusetts Educational Assessment Program

The National Assessment of Educational Progress

The National Assessment of Educational Progress (NAEP) began in 1969 as a large-scale data collection effort to assess the state and progress of education in the United States. Since then, NAEP has assessed students many times in several different content areas. In the spring of 1986, NAEP assessed 9-, 13-, and 17-year olds, as well as students in grades 3, 7, and 11, in reading, mathematics, science, and computers. The full battery of NAEP questions, including both cognitive and questionnaire items (but excluding the subject area of computers) totalled 2,501 questions. These materials were administered to a total of 88,708 students in several hundred schools.

For the 1987-88 school year, NAEP changed the grade levels its assessed to 4, 8, and 12. Reading, writing, civics, geography, and U.S. history were the areas primarily assessed. That assessment involved approximately 130,000 middle, elementary, and high school students from 1,500 schools nationally.

Differences between NAEP and the Massachusetts Assessment

No matter how much effort is put into making two assessments the same so that results will be comparable, differences of sufficient magnitude invariably exist and must be considered when interpreting the data. For example, Chapter 188 mandates that the results of the Massachusetts assessment be made publicly available, school by school, which is a significantly different method of reporting from that done by NAEP. Thus, simple adherence to legislative requirements means that comparisons between NAEP and Massachusetts data are not direct and must be interpreted with some caution. This appendix details the differences between the two assessments, explains what is done to minimize the effect of the differences, and discusses the impact that the differences might have on the results.

Reporting. As noted above, the reporting differences between NAEP and what is required for the Massachusetts assessment by Chapter 188 are substantial. NAEP releases test results for the United States as a whole, thereby guaranteeing anonymity to all participating schools, districts, and even states. In contrast, all participants in the Massachusetts assessment know, in advance of the testing, that results of that assessment will be publicly reported at the school and district level. The motivations of the two groups when tested, therefore, are considerably different. It can only be assumed that Massachusetts students and administrators are motivated to do their best, and, as a consequence, will score higher than the national sample even if the achievement levels of Massachusetts students are no higher than their counterparts nationwide.

Population tested. Chapter 188 mandates that all students at the applicable grade levels, with limited exemptions, will be tested in the Massachusetts assessment. These rules are different

from those used by NAEP; thus the population of students eligible to be tested in Massachusetts is somewhat different from the national sample. NAEP's rules call for the testing of all students except those "deemed unassessable by the school authorities."

One impact of the effect of local reporting is clear. NAEP tests are administered by a person who is sent to the school with a list of students to be tested. That administrator reviews the list with the principal of the school, and they decide jointly about who should be tested and who should not. In this instance, the principal has just two motivations: to adhere to the guidelines for NAEP, and to serve the best interests of students who might be adversely affected by participating in the testing.

In contrast, the Massachusetts assessment has stricter guidelines on who will be tested (that is, all students are tested unless specific criteria are met to exclude them), but local administrators have greater motivation to exclude students who are not likely to score well on the tests. In Massachusetts, a high percentage of students who are not special needs or limited-English-proficient are tested. It could be argued that Massachusetts performance is lowered because of that; however, it is not known how many special needs and limited-English-proficient students who could have been tested, and therefore would have been tested for NAEP, were excluded from the Massachusetts testing.

A second distinction between the groups tested in the two assessments that does allow for some margin of difference is the fact that the Massachusetts assessment is census testing (that is, every student in every school is to be tested), while NAEP uses a sampling design (that is, a sample of students in a sample of schools are tested). NAEP uses a complex weighting system to estimate national performance based upon the sample they have drawn, but there always remains some element of doubt about the actual implementation in the field of sampling designs.

Test format. The NAEP budget for two years of testing is $8 million; the Massachusetts assessment, which involves the testing of more students in more schools, had a budget of $1.3 million. Obviously, some changes had to be made in the Massachusetts assessment to keep costs down. Some of these changes probably have some impact on the direct comparability of the results.

A major difference is in the test materials themselves. Because of the volume of the Massachusetts assessment, it is important to limit the number of pages included in each booklet. Thus, while NAEP typically places three or fewer test questions on one page of text, the Massachusetts assessment test booklets usually are double-columned, and contain three to six questions per page. Also, while the Massachusetts assessment tries to use pictures and charts exactly as used by NAEP, sizes occasionally were changed and the clarity of reproduction was not the same. NAEP asks students to respond to questions directly in the test booklet, while at grades 8 and 12, Massachusetts students record their answers on separate answer sheets. In contrast to the concerns raised in the previous section, there is considerable literature about the effects these changes would have made. In every case, they put Massachusetts students at a disadvantage, and thus tends to lower the performance of those students relative to their counterparts nationwide.

Use of trained administrators. NAEP sends staff to every school participating in the testing to oversee test administration. Massachusetts mails all their materials to schools, and test administration is a local effort. Training workshops are conducted regionally in the state to properly

prepare staff for the administration of the tests, and manuals detailing the administration procedures are included with all testing materials, but there is no requirement that the workshops be attended or the manuals read. The impact of this difference on test results is unknown.

Timed tests. Some testing procedures that are appropriate for NAEP, given the intent of that program, must be modified for the Massachusetts assessment. For example, the NAEP tests are timed. Each student is given a short period of time to answer a series of attitudinal and test questions. It was not unusual in the 1986 assessment for forms to be completed by fewer than half the students taking them. Given the direct relationship between testing and curriculum evaluation that is an outcome of the Massachusetts assessment, it was considered unacceptable to limit the time students had to answer the questions, and therefore the Massachusetts tests were untimed.

Use of "I don't know." Another difference between the two programs revolves around guessing and related issues. Consistent with its mission, NAEP frequently uses the distractor "I don't know" in its multiple-choice questions. While that answer option makes great sense for NAEP, it provides a confusing message to students taking the Massachusetts assessment. Because results are reported locally in Massachusetts, and because answering "I don't know" clearly is an incorrect answer, Massachusetts students should never select that option when it is presented. While some do, it can only be assumed that administrators must discourage them from doing so. Also, NAEP provides no instructions to students about what to do if they don't know the answer to a question. Again, given local reporting requirements of the Massachusetts assessment, it is necessary to provide some direction about guessing so that tests are administered uniformly statewide. The Massachusetts assessment directions therefore instruct students to "make your best guess," while the NAEP directions do not.

Attempts to Make the Data As Comparable As Possible

Given all the issues raised above about the comparability of the data from the two assessments, it was necessary to plan carefully to make the data as comparable as possible. This section describes those plans and the changes that were made in them as the assessment progressed.

As noted above, certain administrative changes had to be made for the Massachusetts assessment in order to meet the administrative requirements of the Massachusetts testing program. For example, NAEP uses several questions that require a short written response ("open-ended questions") rather than being multiple-choice. The administrative restrictions on the Massachusetts tests do not permit switching item types within a block. Rather than discarding the open-ended questions, Massachusetts converted the open-ended questions to multiple-choice for the 1986 assessment. No NAEP-comparable data could be obtained, of course, when an item was changed significantly; by including the item, however, the context in which students took the multiple-choice questions around those open-ended questions remained as similar as possible.

A major substudy undertaken during the 1986 assessment influenced the construction of the Massachusetts tests in 1988 and 1990. The difference between NAEP comparisons between items presented in intact NAEP blocks and those scattered among other questions in the test was found to be insignificant. This finding led to the construction of tests in the latter two years that included no

intact blocks, but rather had NAEP items included at random. Consequently, the practice of converting and including NAEP open-ended questions became no longer necessary.

Another difference between NAEP and the Massachusetts assessment is the use by Massachusetts of separate answer sheets on which students record their answers. So that students make as few tracking errors as possible, the Massachusetts answer sheets are custom designed for each form so that the correct number of questions and the correct number of options for each question are printed on the answer sheet for each student. Since each answer sheet matches only one booklet, the procedure also helps to assure that the number of the form the student takes is correctly recorded on the answer sheet.

The final modifications in the plans to maximize the comparability of the data came after receiving the 1986 NAEP data tapes. Through those, it was confirmed that many students did not finish the NAEP blocks, and thus, students in Massachusetts had a major advantage in taking the test. In order to minimize this problem, only test questions which had been reached by at least 95 percent of the NAEP sample were used in making comparisons. This resulted in discarding a great deal of information that was available, but it was felt that any attempts to interpret results on questions that were not reached by large numbers of students would be fraught with error. As noted above, this was a problem only for the 1986 data. In 1988, NAEP changed its administration procedures so that most of the students completed its tests.

MASSACHUSETTS DEPARTMENT OF EDUCATION

REGIONAL CENTERS

Paul Burnim
Greater Springfield Regional Center
Macek Drive
Chicopee, MA 01013
413/594-8511

Laurie Slobody
Central Mass. Regional Center
Beaman Street, Route 140
West Boylston, MA 01853
617/835-6266

Amy Sosman
Northeast Regional Center
Jefferson Office Park
790 Turnpike St.
North Andover, MA 01845
508/689-0164

Pat O'Brien
Southeast Regional Center
P.O. Box 29
Middleboro, MA 02346
617/947-1231

Related publications about the Massachusetts Educational Assessment Program:

- *The Massachusetts Educational Assessment Program: 1988 Statewide Summary*
- *Using the School Report*
- *Reading and Thinking: A New Framework for Comprehension*
- *Science in the Elementary Schools*
- *Moving Geometry From the Back of the Book: A Report on Geometry and Measurement in the 1986 Assessment*
- *Social Studies Report*
- *On Their Own: Student Response to Open-Ended Tests in Reading*
- *On Their Own: Student Response to Open-Ended Tests in Mathematics*
- *On Their Own: Student Response to Open-Ended Tests in Science*
- *On Their Own: Student Response to Open-Ended Tests in Social Studies*

Advanced Systems IN MEASUREMENT & EVALUATION, INC.
P.O. BOX 1217 • 171 WATSON ROAD, DOVER, NEW HAMPSHIRE 03820 • (603) 749-9102